SOPER: Discovering the influence of fashion and the many faces of User from Session logs using Stick Breaking Process

Lucky Dhakad
Flipkart
lucky.d@flipkart.com

Mrinal Das*
Indian Institute of Technology, Palakkad
mrinal@iitpkd.ac.in

Chiranjib Bhattacharyya
Indian Institute of Science
chiru@csa.iisc.ernet.in

Samik Datta
Flipkart
samik.datta@flipkart.com

Mihir Kale
Flipkart
mihirkale.s@flipkart.com

Vivek Mehta
vivekm@gmail.com

ABSTRACT

Recommending *lifestyle* articles is of immediate interest to the e-commerce industry and is beginning to attract research attention. Often followed strategies, such as recommending popular items are inadequate for this vertical because of two reasons. Firstly, users have their own personal preference over items, referred to as personal styles, which lead to the long-tail phenomenon. Secondly, each user displays *multiple personas*, each persona has a preference over items which could be dictated by a particular occasion, e.g. dressing for a party would be different from dressing to go to office. Recommendation in this vertical is crucially dependent on discovering styles for each of the multiple personas. There is no literature which addresses this problem.

We posit a generative model which describes each user by a Simplex Over PERsona, SOPER, where a persona is described as the individuals preferences over prevailing styles modelled as topics over items. The choice of simplex and the long-tail nature necessitates the use of stick-breaking process. The main technical contribution is an efficient collapsed Gibbs sampling based algorithm for solving the attendant inference problem.

Trained on large-scale interaction logs spanning more than half-a-million sessions collected from an e-commerce portal, SOPER outperforms previous baselines such as [9] by a large margin of 35% in identifying persona. Consequently it outperforms several competitive baselines comprehensively on the task of recommending from a catalogue of roughly 150 thousand lifestyle articles, by improving the recommendation quality as measured by AUC by a staggering 12.23%, in addition to aiding the interpretability of uncovered personal and fashionable styles thus advancing our precise understanding of the underlying phenomena.

CCS CONCEPTS

•Information systems →Data mining;

KEYWORDS

fashion, lifestyle, topic models, Bayesian nonparametrics, stick-breaking process

*This work was done while at University of Massachusetts Amherst.

CIKM'17 , Singapore, Singapore

DOI: 10.1145/3132847.3133007

1 INTRODUCTION

With the advent of Internet, e-commerce portals have started seeing significant activity in the lifestyle segment. Products such as *apparels, fashion accessories* are often classified as lifestyle articles. Recommending such products are considered to be difficult and is beginning to gain attention in academic community [9, 18].

The difficulty in recommending lifestyle products arises primarily due to certain unique characteristics. Firstly, sales of lifestyle products [9] exhibit a long-tail behaviour where large number of products are sold only a few times. This can be attributed to the fact that lifestyle vertical offers substantially larger number of choices in every category of products when compared to other categories such as *consumer electronics*. This allows each user to demonstrate their individual preferences in choosing a product, which often results in items being sold in small numbers. Secondly, it is also to be noted that user choices are not solely dependent on their individuality but also depends on prevailing popular trends of the day, often called fashion. Thirdly, users display multiple personas while making clothing choices. Persona can be understood as the fact that certain combinations of products are well suited for certain occasions.

Understanding user preferences and the ability to elicit fashionable styles from such long-tail phenomena remains a challenging open problem which may lead to more effective recommendation. Interestingly, it has been observed that, profit in the long-tail is significantly more in e-commerce than physical stores [14]. Moreover, recommending products at long-tail catering to users style provides a one-stop-shopping feeling and increases overall customer satisfaction [1, 3].

In this paper we address some of the aforementioned open problems. There is no accepted definition of *style*, but [6, 18] suggests that a combination of items preferred by the users can be called as styles. This idea has been further refined in a recent paper [9] where style is defined as a probability distribution over items which paves the way for using tools such as Latent Dirichlet allocation (LDA) [2]. It partially addresses the problem of eliciting fashionable styles, however it misses some of the styles which are popular only in a sub-community, often called Alt-fashion[1]. But the biggest drawback is there is no provision of finding personas.

The aim of this paper is to study the problem of discovering personas of individuals and make appropriate recommendations for lifestyle articles. In particular, we attempt to answer the following questions: (a.) Can we identify *persona* of a user from collective session logs? (b.) Often, popular styles of the day are called *fashion*,

[1] https://en.wikipedia.org/wiki/Alternative_fashion

can we detect personas of users influenced by fashion ? (c.) Can we detect consistent styles ? (d.) Can we utilize the discovered personas to develop better recommendation algorithm ?Keeping this question in mind we make the following contributions

Contributions.

- The main contribution of this paper is a novel generative model which views users preference as a simplex over personas, SOPER[2]. SOPER explicitly tries to capture the influence of styles on the persona of each user and is naturally suited for handling the long-tail phenomena unique to the lifestyle vertical. In keeping with Bayesian nonparametric setup SOPER uses stick-breaking process (SBP) [11] as a prior over the simplices.
- The elegance of SBP brings in the challenge in inference mechanism which is non-standard in topic models dealing with Drichlet distributions. However, we could derive an efficient collapsed Gibbs sampling inference, that requires to sample only one additional set of random variables compared to LDA.
- Based on SOPER, we develop a recommendation algorithm that utilizes the simplex over personas. We propose to recommend items by identifying users with similar persona as identified by SOPER. The recommendations comprehensively outperform the state-of-the-art LDA-based baseline [9] and several other competitive baselines.
- SOPER can be used to uncover various insights from the data which are otherwise hard to discover. Among other things, we are able to identify users who lack individuality when it comes to purchasing lifestyle articles and can be said to be affected by fashion of the day. Moreover, the styles discovered demonstrates considerable interpretability and purity.

2 ANALYSIS OF LIFESTYLE CLICK-LOGS AND PROBLEM FORMULATION

In this section, we describe data sampling and filtering steps and narrate salient observations on the existence of multiple buyer personas.

2.1 Dataset description

Our datasets comprise of samples drawn from anonymised *click-logs* harvested at an e-commerce portal that serves $\sim$ 100 million registered users in India. The click-logs comprise of user-item interactions aggregated across several facets of engagement, e.g. product search, recommendation, and browsing product listings. We randomly sample $\sim 10K$ users who had interacted with several lifestyle categories - spanning apparels, fashion accessories, and, beauty and personal care products - over a span of $\sim$ 18 months. Further, we retain only their high-intent interactions - additions to wish-lists and shopping carts, and purchases. Finally, we prune away users containing unusually high number of purchases(>1000) and sessions(>700), and also prune away products with either very high(>1000) or very low(<5) purchase frequencies.

[2]One could expand the acronym as *Simplex Over PERsona*

Table 1: Descriptive statistics for DM and DW .

Dataset	U	$\lvert\mathcal{I}\rvert$	$\lvert\mathcal{L}\rvert$
DM	$10K$	$\sim 150K$	$\sim 1.5M$
DW	$10K$	$\sim 150K$	$\sim 2.5M$

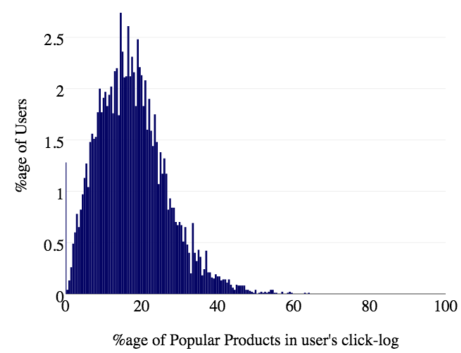

Figure 1: x-axis denotes percentage of popular products present in $\mathcal{L}_u$, and y-axis denotes the percentage of users belonging to the corresponding bucket. We see a mode near 20%, as well as a skew towards the left-tail indicating that most users interact with less-popular items.

To study gender-specific idiosyncrasies, the click-logs are further classified into two datasets, **DW** and **DM** , comprising of interactions with lifestyle categories that cater to the feminine and the masculine genders, respectively.

We will use the following notation throughout the paper. Let $\mathcal{L} = \cup_{u \in \{1,\ldots,U\}} \mathcal{L}_u$ denote a collection of user-specific click-logs, where the anonymised users are represented with unique identifiers u. Each of these click-logs, $\mathcal{L}_u$, records user interactions of the form $\{e_t \mid t \in \mathcal{T}_u\}$ interpreted as follows: the user u had interacted with the item, $e.i \in \mathcal{I}$, at time t during her tenure, $\mathcal{T}_u$. The item attributes, as obtained from the lifestyle product catalogue, is generically represented as $i.a \in \mathcal{A}$, where $\mathcal{A}$ refers to the universe of item attributes spanning colour, fabric, price, and the occasion it is suitable for.

We further partition each of these click-logs, based on the event time-stamps, into *sessions*, $\{\mathcal{L}_{u,s} \mid s \in \mathcal{S}_u\}$, where every new session is assumed to be initiated after 30 minutes of inactivity. In other words, for every $e_t \in \mathcal{L}_{u,s}$ and $e_{t'} \in \mathcal{L}_{u,s'}$, where $s \neq s'$, $|e_t - e_{t'}| \geq 30m$. The scale of these two datasets are summarised in Table 1.

2.2 Evidence of multiple buyer personas

Unlike the electronics and home appliances categories, where buyers' choices tend to align with population-wide choices, lifestyle buyers demonstrate a great degree of diversity in terms of their choices as evident from figure 1. The top-1% items in lifestyle appeal to less than 5% of the buyers, and more than 85% of the buyers purchased items that fall in the bottom 5% of popularity, demonstrating the long tail of choices.

Furthermore, we observe that roughly 55% of the buyers had shopped for 4 or more occasions in **DM** , whereas it is $\sim$ 75% in the **DW** dataset. These observations, together, has motivated us to model an user as simplex over several *personas*.

2.3 Problem description: Modelling buyer personas

To explain the aforementioned diversity of choices and the presence of diverse shopping occasions within $\mathcal{L}_u, \forall u \in [U]$, we postulate the existence of multiple buyer personas for each user. We investigate the problem of learning these personas from the click-logs, $\mathcal{L}$, and further investigate their conformity with the popular fashion.

Existing literature around lifestyle item recommendation in e-commerce (see [9] and references therein) is sparse and have not yet attempted the problem of discovering personas. What makes the problem particularly challenging is the high degree of variability across users; with some users possessing only one persona, while the rest clearly manifesting a multitude of them.

3 SOPER: SIMPLEX OVER PERSONAS

In this section, we will develop the proposed model SOPER. First we will discuss relevant background and notations followed by the principal approach. Then we will describe details of SOPER, associated inference algorithm, comparison with the state of the art methods, and the recommendation algorithm based on SOPER.

3.1 Notation and background

We define notations here, which will be used throughout this section. δ_x denotes an atomic distribution where x is called as atom. $P = \sum_i w_i \delta_{x_i}$ denotes that $p(x_i|P) = w_i$. (x_i) denotes an ordered set of variables with index i. $\{x_i\}$ denotes a set of same class of variables. $[k]$ denotes integers 0 to k. 1_I denotes a vector of length I, where each element is one.

LDA[2] is a probabilistic generative model that has been used earlier by [9] to model lifestyle data. LDA uses a single distribution over styles for each user, thus it uses one persona for a user. Styles are shared among the users, but proportions over them vary across the users. For each activity, LDA first samples a style given the persona of the user, and given the style it samples an item. Thus, LDA can model personal preferences but cannot separate out multiple personas.

Dirichlet Process (DP) [5] is a Bayesian NonParametric (BNP) prior that can be used to automatically select model complexity given observations. DP is highly explored but is not suitable in many cases, where SBP can be useful.

Stick-breaking process (SBP) can be defined as follows. Any a.s. discrete probability measure $\mathcal{P}$ is a stick-breaking process (SBP) [11] if it can be represented as

$$\begin{aligned} \mathcal{P} = \textstyle\sum_{i=1}^{\infty} w_i \delta_{x_i}, w_1 = v_1,\ w_j = v_j \prod_{l=1}^{j-1}(1 - v_l) \\ a_j, b_j > 0,\ v_j \sim Beta(a_j, b_j),\ \ x_j \sim \mathcal{H} \end{aligned} \tag{1}$$

$\mathcal{H}$ is a *diffuse measure* over a measurable space $(\Omega, \mathcal{B})$ and $\{a_j, b_j\}$ are set of parameters. The construction of (w_i) assures that $\sum_{i=1}^{\infty} w_i = 1$ which makes $\mathcal{P}$ a probability measure over countably infinite number of atoms. Stick-breaking representation of DP is a special case of SBP under suitable choice of parameters.

3.2 Approach

We first describe our approach intuitively here, discussing the limitations of the state-of-the-art methods.

Existence of multiple personas is natural in human beings, where a persona reflects a state of the mind or preference[3]. Any individual can go for online shopping with various personas in different times. For example, a *school teacher* who is also a *mother*, a *wife*, and a *daughter* will have at least five personas. When she is shopping for the *annual festival at the school*, she will be looking at very different items than what she will be looking as a daughter to buy a gift for her *mother's birthday*.

The major limitation of the existing methods is that, they either (i) split the information corresponding to a user into multiple parts based on sessions, and model each of them independently, or (ii) model the entire information corresponding to a user using a single persona. Both the options are suboptimal due to the following reasons. In the first case, models will fail to utilize the fact that a user can wear the same persona in multiple sessions, so modeling them together can benefit in learning. However, the challenge is that, it is a priori hard to know which sessions cater to the same persona. We need an automated way to learn that implicitly. In the second case where models use a single persona, they clearly miss the fact that a user has multiple quite different preferences. Considering our example above, representing a user's interest as a mix of *story books*, *ethnic wears*, and *kitchen accessories* cater to a very gross level picture and fails to sufficiently specialize the model to the specific needs of a person depending on her current state of the mind at the time of shopping. One immediate impact of these limitations is that, such models fail to recommend items which are in the long-tail of the users' preferences.

In this paper, we propose to model a user using multiple personas, which is the key component in the proposed method SOPER. In the probabilistic generative model framework of SOPER, we use a distribution over multiple personas for each user, and when a person logs into a shopping session, SOPER samples a suitable persona first.

Once, we get a persona, we get a preference for the user. For example, for the *mother* persona in our example, we get a preference for *children section*, and she may look for items suitable for her children which can be from clothes, sports, toys or books sections. In the probabilistic framework of SOPER, we define a persona as a distribution over styles, where style is a coherent collection of items, where coherence is defined in terms of occasion, purpose, category etc [9]. Thus there are styles specific to the items in the children section for clothes, sports, toys or books. Styles are common across all the users, and there are finite number of them. But different users have different proportions over them, which are the personas for the users.In the following subsections we will describe the details of the proposed method SOPER.

[3] https://en.wikipedia.org/wiki/Persona

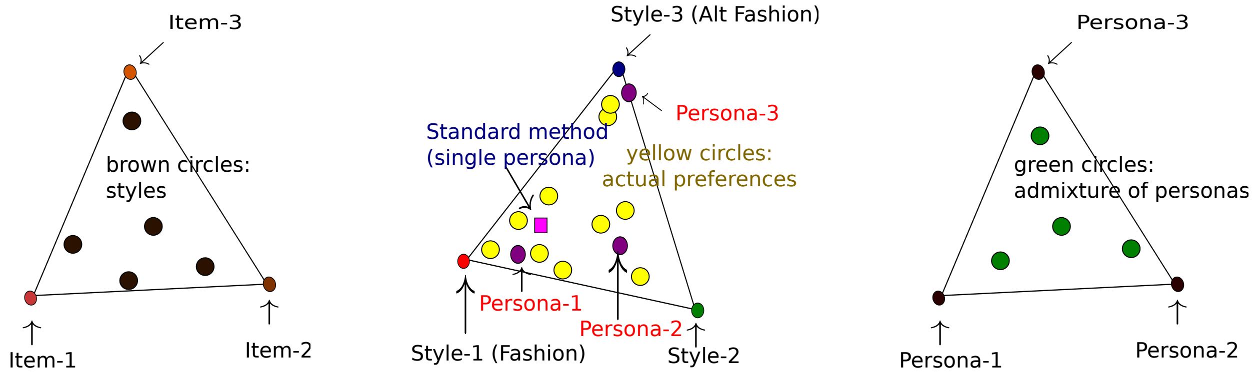

Figure 2: Illustration of the main concept in SOPER. The left simplex shows styles as distribution over items. The middle simplex shows persona as distribution over styles. The right simplex shows user as admixture of personas. Major contribution in SOPER is the middle and the right simplices. In the middle simplex, there are three styles at the vertices. Style-3 is alt-fashion (less popular). Yellow circles denote preference of a user in various sessions. Standard methods using one persona tries to emphasize the fashionable style. SOPER using multiple personas (three here) can model users preference better by covering a larger probability space. For example, persona-3 can model alt-fashion for this user effectively. The right simplex is modeled using SBP, that allows to model users multiple personas from activity logs. Figure 3 gives the generative model description of SOPER that implements this concept.

3.3 Style and persona

Style. Recall that, style is a distribution over items [9]. We denote styles using $\{\beta_s,\ s = 1, 2, \ldots, S\}$, considering S styles, and define each style as

$$\beta_s = \sum_{i=1}^{I} \beta_{si}\delta_i, \quad \text{s.t. } (\beta_{s1}, \ldots, \beta_{sI}) \sim Dirichlet(\eta 1_I), \tag{2}$$

where η is a scalar hyper-parameter. Dirichlet is a suitable distribution to model such random variables, so that $0 \leq \beta_{si} \leq 1$, and $\sum_{i=1}^{I} \beta_{si} = 1$. It is appropriate to mention here that, styles can be fashionable if they have high weight in terms of β_{si} over popular items in the dataset. Otherwise, if a style has high weight over unpopular items, they correspond to alt-fashion.

Persona. A persona of a user relates to the preference over styles keeping in mind some impending ocassion or a role. Topic modeling literature provides a natural definition to such objects.

$$\gamma_u = \sum_{s=1}^{S} \theta_{us}\delta_{\beta_s}, \quad \text{s.t. } (\theta_{u1}, \ldots, \theta_{uS}) \sim Dirichlet(\alpha 1_S) \tag{3}$$

Again, the Dirichlet prior over (θ_{us}) ensures that $0 \leq \theta_{us} \leq 1$, and $\sum_{s=1}^{S} \theta_{us} = 1$. θ_{us} indicates the weight of the persona γ_u on the global style β_s. If a persona shows a strong preference on one of the styles, e.g. $\theta_{us} \geq 0.9$, then the style indexed by s could be understood as strongly influencing the user.

State of the art method by [9] comprises of the components described above. One key limitation of their work is to use single persona for each user, and thus they can model only one preference over the styles. Modeling multiple personas of users using their activity logs is non-trivial, and there is no immediate work that addresses this problem (discussed more in Section 3.8), which is the main contribution of the paper.

3.4 User is an admixture of her personas

Experimental data suggests that in most sessions only one of the personas are active but there are sessions in which there are multiple personas in play. If in each session only one of the personas were active one could have used a simple probabilistic mixture model. Keeping in mind the empirical evidence we posit that simple mixture maybe inadequate and a more appropriate model of user could be an admixture of her personas.

Simplex over personas. This hypothesis is modelled as follows, we use a set of J personas $\{\gamma_{uj}\}_{j=1}^{J}$, and for each user u, we define a simplex over personas as follows.

$$\Gamma_u = \sum_{j=1}^{J} \rho_{uj}\delta_{\gamma_{uj}},\ \gamma_{uj} = \sum_{s=1}^{S} \theta_{ujs}\delta_{\beta_s} \tag{4}$$

where $\{\rho_{uj}\}$ should be defined in such a way that $0 \leq \rho_{uj} \leq 1$, and $\sum_{j=1}^{J} \rho_{uj} = 1$. Figure 2 illustrates the concept of SOPER.

Discovering influence of fashion through personas. ρ_{uj} defines a weight over the jth persona of the uth user. On the other hand, for the jth persona γ_{uj}, θ_{ujs} defines the weight over the sth style, and β_{si} defines the weight over the ith item. To understand the influence of fashion on users, we define following variables:

$$\phi_{ujs} = \rho_{uj}\theta_{ujs},\ \psi_{uji} = \textstyle\sum_{s=1}^{S} \phi_{ujs}\beta_{si}. \tag{5}$$

Thus, ϕ_{ujs} denotes the influence of the sth style on the uth user reflected in the jth persona. If sth style has high probability over the popular items, then ϕ_{ujs} corresponds to the influence of fashionable style to user u, otherwise ϕ_{ujs} corresponds to the individualistic taste of the user.

ψ_{uji} gives the probability of the ith item corresponding to the jth persona of the uth user. By looking into the popularity of items in the dataset and comparing with ψ_{uji}, we can categorize the jth

- For $s = 1, 2, \ldots, S$
  - Sample styles $\beta_s \sim Dirichlet(\alpha 1_S)$
- For each user $u = 1, 2, \ldots, U$
  - Sample simplex over persona $\Gamma_u \sim SBP(a_1, a_2)$
  - For each activity session $a = 1, 2, \ldots, A_u$
    - For each item $i = 1, 2, \ldots, I_{ua}$
      - Sample persona $\mu_{uai} \sim \Gamma_u$
      - Sample style $\nu_{uai} \sim \mu_{uai}$
      - Sample product item $x_{uai} \sim \nu_{uai}$

Figure 3: Model description of SOPER. One key component is the simplex over personas for each user Γ_u with SBP as a prior. Figure 2 illustrates the concept.

persona of the uth user as fashionable or individualistic. It is found in our experiments that, a large number of users have individualistic as well as fashionable personas. LDA completely misses this fact, and is forced to label users as either fashionable or individualistic.

In order to model personas, we advocate a Bayesian nonparametric (BNP) route and hence we would require a suitable prior over the simplex governing the admixture. We are not aware of any suitable prior which could work with the overall model. We posit an *SBP* prior over this model which we describe next.

3.5 SBP as a prior for simplex over personas

Using SBP as a prior we model the simplex over personas as $\Gamma_u \sim SBP(a_1, a_2)$. Following Eq. (1), we describe it as follows:

$$\Gamma_u = \textstyle\sum_{j=1}^{J} \rho_{uj} \delta_{\gamma_{uj}},$$

$$\rho_{u1} = v_{u1}, \rho_{uj} = v_{uj} \textstyle\prod_{l=1}^{j-1} (1 - v_{ul})\ j \in [J-1], v_{uJ} = 1, \quad (6)$$

where $v_{uj} \sim Beta(a_1, a_2)$. Figure 3 describes the generative model of SOPER.

Advantage of SBP. The challenge in modeling multiple personas is that, Γ_u will be very different across the users, as some users have high variability in preferences leading to a larger number of personas, whereas some users have low variability leading to a low number of personas. Therefore, the effective size of the simplex over personas Γ_u will be very different across users. BNP priors have been found useful in such cases.

DP is a common choice in BNP, and is effective in general mixture models. However, DP has two limitations: (i) it fails to model high variability within a document due to its high rich getting richer effect, that is one atom tends to get more attention, (ii) DP does not allow sharing of atoms across measures, that is personas can not be shared across different activity sessions, and one needs to resort to a truncated ad-hoc version which is a very special case of SBP [11]. On the other hand, SBP allows sharing of atoms, and also allows to model high variability within a document.

Collapsed Gibbs sampling inferences are generally preferred in topic models due to its simplicity and efficiency. However, SBP is highly un-explored in topic models, and inference mechanism is non-standard. We will need to marginalize out the weights over personas $\{\rho_{uj}\}$ (Eq. 6), which is the main challenge in deriving a collapsed Gibbs sampling inference mechanism of SOPER. We discuss our approach below.

3.6 Inference procedure for SOPER

The generative model gives the modeling principle, and through inference given the observed user activities $\{x_{uai}\}$ we infer the variables corresponding to style, persona.

The usual trick for collapsed Gibbs sampling inference is change of variables. We will follow that, and simplify the inference procedure to marginalize out all the continuous random variables and work with discrete random variables. Note that, observations x_{uai}s are discrete variables, and using conjugacy with Dirichlet distribution we can collapse styles $\{\beta_s\}$. Let $z_{uai} = s$ iff $\nu_{uai} = \beta_s$, and $b_{uai} = j$ iff $\mu_{uai} = \gamma_{uj}$. By $\mathbf{z}$ and $\mathbf{b}$ we will denote the set of all z and b variables. The main challenge in the inference is to infer $\mathbf{b}$, which also denotes the additional set of variables over LDA.

3.6.1 Inference mechanism for SBP prior. The atoms i.e. persona $(\gamma_{u1}, \gamma_{u2}, \ldots, \gamma_{uJ})$ are independent of the weights; so we can decouple the inference for weights over persona and inference for persona. The inference due to persona will be done using the conjugacy between Dirichlet and multinomial distribution. For the weights, we will utilize the relationship between SBP and *generalized Dirichlet distribution*(GDD) [4]. The weights over atoms in SBP $(\rho_{u1}, \rho_{u2}, \ldots, \rho_{uJ})$ are distributed as GDD. Interestingly, GDD is also conjugate to multinomial distribution similar to Dirichlet distribution. That will help us to integrate out the weights. Following, GDD the density of $\rho_u = (\rho_{uj})$ is:

$$f_{\rho_u} = \prod_{j=1}^{J-2} \frac{\rho_{uj}^{a_1-1} (1 - \sum_{l=1}^{j} \rho_{ul})^{-\kappa_l}}{B(a_1, a_2)} \quad (7)$$

where $B(a_1, a_2) = \frac{\Gamma(a_1)\Gamma(a_2)}{\Gamma(a_1+a_2)}$. $\kappa_l = -a_1$ for $l = 1, 2, \ldots, J-2$ and $\kappa_{J-1} = a_2 - 1$. Note that, $\rho_{uJ} = 1 - \sum_{l=1}^{J-1} \rho_{ul}$.

Now using the conjugacy between GDD and multinomial we integrate out ρs and vs. Following the definition of b_{uai}, we can say that $b_{uai} \sim multinomial(\rho_u)$, and $\rho_u \sim \mathcal{GD}_{J-1}(a_1, a_2)$, then the posterior distribution of ρ_u given (b_{uai})s is again a GDD with density

$$\mathcal{GD}_{J-1}(\bar{a}_{11}, \ldots, \bar{a}_{1J-1}, \bar{a}_{21}, \ldots, \bar{a}_{2J-1})$$

, where $\bar{a}_{1j} = a_1 + n_{uaj}^{-uai}, \bar{a}_{2j} = a_2 + \sum_{l=j+1}^{J} n_{ual}^{-uai}$. where n_{uaj} is the count number of times $b_{ual} = j$ and n_{uaj}^{-uai} is without counting for assignment of b_{uai}.

Thus we compute conditional $p(b_{uai} = j | b^{-uai})$, for $j < J$ as

$$\frac{a_1 + n_{uaj}^{-usi}}{a_1 + a_2 + \sum_{r=j}^{J} n_{uar}^{-uai}} \prod_{l<j} \frac{a_2 + \sum_{s=l+1}^{J} n_{ual}^{-uai}}{a_1 + a_2 + \sum_{s=l}^{J} n_{ual}^{-uai}}$$

and $p(b_{uai} = J | b^{-uai}) = 1 - \sum_{l=1}^{J-1} p(b_{uai} = l | b^{-uai})$.

3.6.2 *Sampling persona.* The posterior probability of selecting persona can be found to be as below:

$$\begin{aligned} & p(b_{uai} = j|\mathbf{b}^{-uai}, \mathbf{z}) \\ \propto \quad & p(z_{uai}|b_{uai} = j, \mathbf{z}^{-uai})p(b_{uai} = j|\mathbf{b}^{-uai}) \\ = \quad & \frac{\alpha + n^{-uai}_{ujz_{uai}}}{K\alpha + n^{-uai}_{uj}} \; p(b_{uai} = j|\mathbf{b}^{-uai}) \end{aligned} \tag{8}$$

where $n^{-uai}_{ujz_{uai}}$ and n^{-uai}_{uj} are the counts of number of times persona j is used with style given by z_{uai} and all styles respectively, excluding the assignment of b_{uai}.

3.6.3 *Sampling style.* The conditional probability of style assignment to item i at session a for user u can be expressed as:

$$\begin{aligned} & p(z_{uai} = s|\mathbf{x}, \mathbf{z}^{-uai}) \\ \propto \quad & p(x_{uai}|z_{uai} = s, \mathbf{z}^{-uai})p(z_{uai} = s|\mathbf{z}^{-uai}) \\ = \quad & \frac{\eta + n^{-uai}_{sx_{uai}}}{I\eta + n^{-uai}_{s}} \; \frac{\alpha + n^{-uai}_{ub_{uai}k}}{K\alpha + n^{-uai}_{ub_{uai}}} \end{aligned} \tag{9}$$

where $n^{-uai}_{sx_{uai}}$ and n^{-uai}_{s} are respectively number of times item x_{uai} is assigned with style s total number of times style s has been used, excluding assignment of z_{uai}.

Equations (8), (9) together form the inference procedure of SOPER.

3.7 SOPER based recommendation algorithm

Armed with the simplices over personas, $\{\Gamma_u\}_{u=1}^{U}$, spanned by $\{\gamma_{uj}\}_{j=1}^{J}$, $\forall u \in [U]$, as learnt with SOPER, we adopt a nearest neighbour based strategy to recommend lifestyle articles to each user by first identifying k most similar users in terms of personas, and then recommending from their interaction history. To enable this, we endow the space of simplices with a distance measure. We begin with representing the user u in terms of the bag of items she has interacted with, f_u, a vector whose ith component is defined by

$$f_{ui} = \frac{n_{u,i}}{|\mathcal{L}_u|} \mid \forall i \in \mathcal{I}, \tag{10}$$

where $n_{u,i}$ is the frequency of the item i in user u's click-log, $\mathcal{L}_u$. Also, we find a projection of f_u on her own persona-simplex to find best fit profile, $\hat{f}_u$, using Equation (11) and putting $v = u$. Next, we use Equation (11) to project $\hat{f}_u$ to the persona-simplex of any other user $v \in [U]$, $\hat{f}_{uv}$

$$\begin{aligned} \underset{\{\alpha_j\}_{j=1}^{J}}{\text{argmin}} \quad & \left\| f_u - \sum_{j \in [J]} \alpha_j \cdot \gamma_{vj} \right\|_2 \\ \text{subject to} \quad & 0 \leq \alpha_j \leq 1, \; \forall j \in [J], \; \sum_{j \in [J]} \alpha_j = 1. \end{aligned} \tag{11}$$

This enables us to compute the persona wise distance between any pairs of users, $(u, v) \in [U] \times [U]$ by computing $\left\|\hat{f}_u - \hat{f}_{uv}\right\|_2$. The resulting k nearest neighbour algorithm operates in the space of simplices thus defined. To find similar users for all the users in the data-set, the complexity of the algorithm is $O(|U|^2 J^3)$.The complexity of projection can be improved from $O(J^3)$ to $O(J \log J)$, since it is a convex problem which includes projection of a vector into simplex of probability.

This algorithm benifits the property of multiple personas found by SOPER in capturing users with similar personas, which is not possible with a single profile of users in state-of-art nearest neighbour methods. Let $N(u, k)$ be the list of k nearest neighbours for user u in sorted order of the distance calculated using the above procedure. The list of items to be recommended to u, $\forall u \in [U]$, is finally generated by arranging the items interacted with by $v \in N(u, k)$ in non-increasing order of frequency and recommending from the top of the sorted list, after excluding the items already purchased by u.

3.8 Related works

LDA has been used earlier to model lifestyle data [9], but LDA fails to model multiple personas. [12] has proposed a topic model in a different context of expertise modeling which is the closest to this work. However, they use one persona to model each document, which is a significant difference, as they assume that authors' role in any document is consistent which is not the case with e-commerce users. Thus, using oner persona per document makes the model by [12] similar to LDA and does not apply here in the problem of modeling users.

Session based models are common in modeling lifestyle data [7, 8, 16, 17]. However, all of them focus on short-term needs of users. This makes it very hard to learn personas of users just looking at the small window of a session. Moreover, they will not work to model item-item and user-user relationships across the entire data and consider high variability in preferences over a large and diverse of items. Another limitation of such models is that, they being based on recurrent neural network (RNN) is able to look into local sequence of items which is suitable at the session level, but fail to model high level co-occurrence statistics, where generative models are found very effective.

There is a different line of work, where users and items are represented in a common space, and compute affinity between users and items to optimize AUC directly [10, 13]. We have used only users' activity logs, without using any information about the items. This family of models do not apply in our case. Nevertheless, it is worth to mention that, recently it has been observed that representing users by multiple vectors can be useful for such models too [19, 20]. [19] considers maximum similarity among the users' vectors with a item's vector, which greedily find maximum possible items for each user, and does not cater to the occasion based persona that we aim here. [20] overcomes this issue by using a projection based method similar to us. However, they fail to model the fact that users' preference is generally coherent within a session, and diverse across the sessions, where similar preference can be observed in different sessions. As a consequence, they lack the ability to perform in the presence of sparsity at the long-tail. The hierarchical Bayesian modeling framework of SOPER allows us to model this phenomenon appropriately which is hard using existing matrix factorization based methods.

4 EMPIRICAL EVALUATION

In this section we will study SOPER empirically. First, we will validate SOPER in terms of the goodness of fit. Next, we will assess the ability of SOPER to discover styles and personas. The work presented in [9] is the closest to our work and is an immediate

Table 2: Comparison on modeling ability in terms of Perplexity (less is better). SOPER is able to increase ability with the increase in number of styles, whereas LDA detoriates.

	DW		DM	
# Styles	20	40	20	40
LDA	13.7	13.8	14.2	14.3
SOPER	**5.3**	**4.1**	**4.3**	**3.1**

baseline. Later we will show application of SOPER to a problem of immense importance in e-commerce – recommendation.

Dataset description. We report results from experiments conducted on two large-scale datasets, **DM** and **DW**, described in detail in section 2.1. We split the datasets into train and test datasets as follows. In order to generate the test dataset for an user $u \in [U]$, we first select one of her sessions, $s \in \mathcal{S}_u$, uniformly at random, and then we set aside one randomly selected item from $\mathcal{L}_{u,s}$ for testing and erase it from the train dataset for u. For each user, the aforementioned procedure is repeated $\frac{|\mathcal{S}_u|}{4}$ times to generate the test dataset. The remainder of $\mathcal{L}_u$ is reserved for training. Note that, for each user, we ensure that no item from the test set is present in the train. Average number of items in test dataset per user for dataset **DW** is 10 and is 8 for **DM**.
We reiterate that the total number of sessions are $\sim 1.5M$ and $\sim 2.5M$ for **DM** and **DW** , respectively, leading to a large scale Bayesian modelling problem. Furthermore, we use categorical attributes of items present in the datasets to analyze and evaluate styles and personas found. These attributes are not used while learning SOPER model.

4.1 Evaluation of modelling ability

The first check one needs to do for evaluating any probabilistic generative model is that, how good it explains some un-seen data. Perplexity is one commonly used metric in this regards. If n_u is the number of events in the test dataset corresponding to user $u \in [U]$, the perplexity can be computed as follows:

$$perplexity = \exp \frac{-(\sum_{u \in [U]} \ln p(w_u))}{\sum_{u=1}^{U} n_u}. \tag{12}$$

Results. The results on **DW** and that on **DM** is tabulated in Table 2. SOPER shows lower perplexity than LDA with good margin on both the data-sets. As we increase the number of styles, it increases complexity of the model, but it is also increasing the ability of the models. Interesting to note that, increasing number of styles in increasing the modeling ability of SOPER and we achieve better perplexity, whereas LDA suffers to utilize the additional complexity.

4.2 Evaluation on discovering styles

Recall that, style is a coherent set of items where coherence is in terms of some attribute. So we verify the coherence or purity in detection of style.

For the design cues, we rely upon [15] and the folklore. From Figure 4, we observe that the style demonstrate purity in that a style rarely contains both party and formal wears, or both light and dark shades for that matter. We also note a few interesting facts: e.g. party wears are typically of dark shades and are expensive, and that the festive wears are expensive and come in light shades. Since the items in each style potentially belong to different categories, we first divide the price ranges for each of those categories into 2 bins based on the median price. From Figure 4, we observe that each style caters to either the high or the low end of the spectrum.

A sample of the styles presented herein can be found in Figure 5. We note that styles are primarily of two kinds: one consisting of items predominantly from the same category, and the other with thematically coherent items from different categories. In particular, style-3 depicts low-priced mixed-shade party wears, whereas style-4 have representatives from lower-body apparels and accessories. It is instructive to note that the corresponding heat maps from Figure 4 are in unison.

The significance of this result is that, SOPER did not have access to any of these item attributes, but SOPER is effective enough to discover coherent styles which can be attributed to the ability to model each user appropriately.

Table 3: Consistent personas found by LDA and SOPER for 20 styles (numbers are percentage). Personas are analysed for different attributes and combination of attributes. SOPER is able to find more consistent personas. Occ. means Occasion and Colr. means Color.

Attribute	Occ.	Colr.	Price & Occ	Price & Colr.	Occ. & Colr	ALL
DW						
LDA	67.4	48.2	66.6	47.6	37.3	36.8
SOPER	**76.8**	**83.0**	**74.7**	**82.9**	**66.7**	**66.6**
DM						
LDA	67.1	58.7	67.0	55.3	41.5	41.4
SOPER	**84.2**	**83.5**	**81.5**	**80.8**	**72.2**	**69.6**

4.3 Evaluation on discovering personas

As we argued in Section 2, and it is also very intuitive that most people have mix of fashionable and individualistic persona. We verify how good our model discovers such personas. We have used quantitative as well as qualitative approaches.

Methodology. As discussed in Section 3.4, to analyse the personas $\{\gamma_{uj}\}_{j=1}^{J}$ of each user u derived from the model, we utilize the quantity $\{\psi_{uji}\}$ defined in Eq. (5). Recall that, high value in (ψ_{uji}) denotes a set of items preferred by the uth user in her jth persona using the sth style. By inspecting top k items in (ψ_{uji}), we can infer several information about the jth persona of the uth user. We validate following properties.

Consistency of discovered personas. We take top 10 products in (ψ_{uji}), and check the attributes of them. If 60% of them belong to the same attribute class, we label the jth persona of the uth user as consistent. In Table 3, we report the result for LDA as well as SOPER. SOPER is able to find more consistent personas.

It can be observed that almost all personas are consistent with *price* attributes, irrespective of the model. It can also be observed

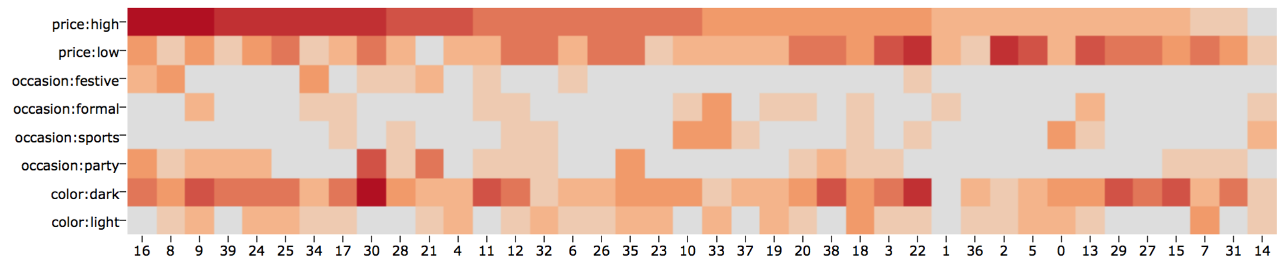

Figure 4: A heat map over item attributes for each of the styles detected by SOPER. Each column represents a single style, denoted by its index. Each row represents an item attribute. For example, top two rows focus on the price, whereas the last two rows depict the colour palette. The styles are sorted such that the price decreases as one reads from left to right.

Figure 5: A sample of styles from DW detected by SOPER. Each column represents one distinct style. Quality of the detected styles is clearly visible here.

Table 4: Percentage of users categorized as fashionable, individualistic, or mix of both fashionable and individualistic. LDA can not discover users of mix category.

Persona	LDA - Fashionable	LDA - Individualistic
SOPER - Mix	40.47	69.36
SOPER - Fashionable	58.27	26.68
SOPER - Individualistic	1.26	3.94

that, even with a subset of attributes SOPER is able to find more consistent personas. Also, *Occasion* and *Color* attribute are not consistent in persona derived from LDA, while SOPER finds consistent personas with *Occasion* and *Color*. It was also observed that SOPER could find distinct consistent personas for more than 50% of users.

Discovering fashionable and individualistic personas. For this evaluation, we have labelled each persona into two kinds: (i) *individualistic*, and (ii) *fashionable*. This labelling is done by checking the top 100 items in (ψ_{uji}). We compute popularity of these 100 items in the dataset and sum them up; if the value is higher than a threshold, then we say that the persona is fashionable otherwise

Table 5: Demographic analysis for fashionable and individualistic users in different categories found by SOPER. SOPER makes it feasible to derive such insights from session logs.

Attribute	Married	Single	15-24	25-34
Mix	0.56	0.44	0.27	0.48
Individualistic	**0.61**	0.39	0.25	**0.48**
Fashionable	0.33	**0.67**	**0.42**	0.45

individualistic. We set the threshold as the average of $\{\gamma_{uj}\}$ across all dimensions, and personas of all users. Using the labels of the personas, we categorize users into one of three categories: (i) all personas are fashionable, (ii) all personas are individualistic, and (iii) mix of both fashionable and individualistic personas.

We notice in Table 4 that, a significant number of users have mix of both fashionable and individualistic personas found by SOPER, which are either labelled as fashionable or individualistic by LDA. We also notice that for 12.6% of users who had shopped for 3 occasions, SOPER accurately retrieves them. The fraction increases to 24% for users who had shopped for 4 distinct occasions.

Demographic study. In Table 5, we analyse sub-populations of users by categorizing them into fashionable, individualistic, and mix of both based on personas inferred by SOPER. We observe that younger users are more fashionable and as age increases their preference becomes more taste inclined, which goes with the intuition because as we grow our taste develops. The marital status of users' shows that users who are single are more fashionable while married user tend to be more individualistic. Such demographic analysis is possible only through SOPER.

Qualitative study. It is intuitive that, more users will have common fashionable personas, but individual personas will vary quite a lot across users. In order to understand this fact, we inspect the personas discovered by SOPER, and we found unison. In Figure 6, we give one instance of discovered personas by SOPER. It could be seen that both user-1 and user-2 have similar fashion choices but their individual taste differs. One user has preferences for low price ethnic wears suitable for casual occasions while other user has preferences of high price ethnic wear suitable for party and festive occasions.

(a) User 1 - Indiv Persona
(b) User 1 - Fashion Persona
(c) User 2 - Indiv Persona
(d) User 2 - Fashion Persona

Figure 6: Example of personas of users tagged using SOPER. Notice that, both users have similar fashionable persona, but they have very different individualistic personas.

4.4 Application on recommendation

After empirically verifying SOPER, we will show one application of SOPER on recommendation.

Evaluation setting. We discuss metrics and baselines for our experiments.

Metrics. We will use AUC, which measures the ability of a recommender system to rank relevant items higher than the irrelevant items. The estimated ranking is evaluated on the test dataset of each user using AUC. Let T_u be the test set for the user u and $\mathcal{L}_u$ be the train set of user u. Following [13], we compute AUC is computed as follows:

$$AUC = \frac{1}{U} \sum_{u \in [U]} \frac{1}{|E(u)|} \sum_{(i,j) \in E(u)} \mathbb{I}(R_i > R_j), \tag{13}$$

where $E(u)$ is set of evaluation pairs defined as

$$E(u) = \{(i, j) \mid i \in T_u \wedge j \notin \mathcal{L}_u \cup T_u\}, \tag{14}$$

and $\mathbb{I}(.)$ is an indicator variable such that it returns 1 if the predicate is true, else returns 0. R_i is the rank of i, assigned by the recommender system under study. A higher AUC value indicates better recommendation quality.

Baselines. LDA based method by [9] is the closest to our work and is an immediate baseline. We also compare the recommendations made by SOPER with several other competitive baselines including variants of the Matrix Factorisation (MF) based approaches.

- **Most Popular (MP) :** This baseline recommends the most popular item to all the users, does not attempt at personalising in any way.
- **SVD-MF:** The SVD-based Matrix Factorisation algorithm.
- **BPR-MF:** Introduced in [13], this algorithm directly optimises a surrogate to AUC.
- **WR-MF:** Weighted Regularised Matrix Factorisation method [10] is a confidence-weighted version of the matrix factorisation suitable for implicit-feedback datasets.

Evaluation by varying number of neighbours. For each data-set, the AUC over test data of LDA and SOPER is compared with varying number of nearest neighbours. From Table 6, it can be seen that AUC values for SOPER increase faster as compared to LDA as the number of nearest neighbours increases. Also, with the increase in number of styles, SOPER can beat LDA with very less number of nearest neighbours. This indicates that, SOPER learns better from users' activity logs than LDA.

Table 6: AUC by varying number of nearest neighbours with number of styles = 40.

k-NN	1	2	5	10	15	20
DW						
LDA	0.54	0.56	0.62	0.67	0.71	0.73
SOPER	**0.60**	**0.64**	**0.7**	**0.77**	**0.80**	**0.82**
DM						
LDA	**0.514**	**0.527**	0.56	0.601	0.631	0.652
SOPER	0.512	0.52	**0.564**	**0.64**	**0.689**	**0.722**

Table 7: AUC by varying number of styles with fixed number of nearest neighbours as 10.

# Styles	20	40	60	80
DW				
MP	0.67	0.67	0.67	0.67
SVD-MF	0.68	0.66	0.64	0.62
WR-MF	**0.69**	0.72	0.74	0.74
BPR-MF	**0.69**	0.72	0.71	0.73
LDA	0.68	0.67	0.67	0.66
SOPER	0.67	**0.77**	**0.83**	**0.84**
DM				
MP	0.65	0.65	0.65	0.65
SVD-MF	0.65	0.64	0.61	0.60
WR-MF	0.68	0.62	0.61	0.66
BPR-MF	0.72	0.69	0.70	0.65
LDA	0.66	0.65	0.64	0.63
SOPER	**0.74**	**0.72**	**0.74**	**0.73**

Evaluation by varying number of styles. The baselines other than LDA do not have a notion of *style*, so we vary number of styles for LDA and SOPER, and vary number of factors for the MF based methods. Table 7 summarizes the results. As the number of factors increase in each baseline, the AUC value increases and shows consistency at high values. SOPER outperforms all other methods consistently except when number of styles is 20. This is understandable because, SOPER learns multiple personas for each user and hence prefers higher number of styles.

Evaluation for fashionable and individualistic users.

Using the method described in Section 4.3, we classify users as fashionable, and individualistic based on their prominent persona. We compare SOPER with LDA on these two sub-populations and report AUC in Table 8. We have kept the number of nearest neighbours fixed to 10. It can be observed that as the number of styles increases LDA and SOPER both improve, but SOPER consistently outperform LDA for fashionable users. On the remaining set of users who display a strong sense of individuality, SOPER beats LDA with a much higher margin. This shows that SOPER captures better *personas* of users.

Table 8: AUC for fashionable, and individualistic users with umber of nearest neighbours as 10.

# Styles	20	40	60	80
DM				
Fashionable				
SOPER	**0.71**	**0.70**	**0.77**	**0.78**
LDA	0.62	0.67	0.68	0.69
Individualistic				
SOPER	**0.65**	**0.63**	**0.64**	**0.65**
LDA	0.59	0.59	0.58	0.58
DW				
Fashionable				
SOPER	**0.67**	**0.84**	**0.89**	**0.91**
LDA	0.67	0.75	0.80	0.82
Individualistic				
SOPER	0.67	**0.76**	**0.82**	**0.833**
LDA	**0.68**	0.66	0.65	0.646

4.5 Discussion

Table 2 asserts that SOPER is not only able to model held-out dataset better than LDA, but also detects consistent styles. These results provide basic validation on the applicability of SOPER on lifestyle dataset. More interestingly, from Figure 4, and Figure 5, we see that styles detected by SOPER correlates with common sense. This is significant because, SOPER did not use item attributes during learning, but while evaluating the performance we see that SOPER is able to learn coherent styles in terms of attributes.

Our findings in Table 4 conform to the fact that many users follow fashion trends whereas there are many users who have more individuality; however most of the users demonstrate mix of both personas. Such observations are beyond the scope of state of the art methods which fail to model personas. One interesting observation from Table 4 is that, LDA tries to label many users as individualistic which is merely due to the fact that, there is a long-tail phenomenon in users logs, that is a large number of users have purchased a large number of unpopular items, which is a common trend in e-commerce. In such cases, fashionable persona of a user becomes rare statistically and LDA finds it hard to detect. Figure 6 illustrates SOPERś ability to uncover multiple personas from the session-logs, by citing two users who possess similar fashionable personas, yet widely different individualistic personas. Table 3 shows that, SOPER is able to find more consistent personas on various attributes and combination of attributes.

Significance of this study is that, SOPER allows us to analyse sub-population of e-commerce users based on various parameters and their preferences; to get more understanding of users through correlation between attributes like gender, occupation etc and their individuality (Table 5). For example, with age people have shown more taste towards individualism. Such analysis gives extra advantage to e-commerce companies to make their business plans.

One key evaluation is through AUC which objectively validates that modeling of personas can model users much better and SOPER is able to find more meaningful nearest neighbors that gives better accuracy in predicting items to recommend.

5 CONCLUSION

Understanding the personas of an indvidual is central to recommending lifestyle products. There has been little work done in this direction. The problem is further complicated by the lack of proper techniques to model personas. Inferring the simplices over personas are not straightforward and require borrowing techniques from recent advances in Machine Learning and Statistics. The model is flexible and can adapt to the change in behaviour of an user.

REFERENCES

[1] C. Anderson. 2006. *The Long Tail: Why the Future of Business Is Selling Less of More.* Hachette Books.

[2] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research* 3 (2003), 993–1022.

[3] E. Brynjolfsson, Y. Hu, and D. Simester. 2011. Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales. *Management Science* 57 (2011), 1373 – 1386. Issue 8.

[4] Robert J. Connor and James E. Mosimann. 1969. Concepts of Independence for Proportions with a Generalization of the Dirichlet Distribution. *J. Amer. Statist. Assoc.* 64, 325 (1969), 194–206. DOI:http://dx.doi.org/10.2307/2283728

[5] Thomas S Ferguson. 1973. A Bayesian analysis of some nonparametric problems. *The annals of statistics* (1973), 209–230.

[6] Jukka Gronow. 1993. Taste and Fashion: The Social Function of Fashion and Style. *Acta Sociologica* 36, 2 (1993), 89–100.

[7] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. 2015. Session-based Recommendations with Recurrent Neural Networks. *CoRR* abs/1511.06939 (2015). http://arxiv.org/abs/1511.06939

[8] Balázs Hidasi, Massimo Quadrana, Alexandros Karatzoglou, and Domonkos Tikk. 2016. Parallel Recurrent Neural Network Architectures for Feature-rich Session-based Recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems (RecSys '16).* ACM, New York, NY, USA, 241–248. DOI:http://dx.doi.org/10.1145/2959100.2959167

[9] Diane J Hu, Rob Hall, and Josh Attenberg. 2014. Style in the long tail: Discovering unique interests with latent variable models in large scale social e-commerce. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining.* ACM, 1640–1649.

[10] Yifan Hu, Yehuda Koren, and Chris Volinsky. 2008. Collaborative filtering for implicit feedback datasets. In *2008 Eighth IEEE International Conference on Data Mining.* Ieee, 263–272.

[11] Hemant Ishwaran and Lancelot F James. 2001. Gibbs sampling methods for stick-breaking priors. *J. Amer. Statist. Assoc.* 96, 453 (2001), 161–173.

[12] D. Mimno and A. McCallum. 2007. Expertise Modeling for Matching Papers with Reviewers. In *KDD.*

[13] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian Personalized Ranking from Implicit Feedback. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence (UAI '09).* AUAI Press, Arlington, Virginia, United States, 452–461. http://dl.acm.org/citation.cfm?id=1795114.1795167

[14] G. Singer and A. Sundararajan. 2010. Recommendation Networks and the Long Tail of Electronic Commerce. *MIS Quarterly* 36 (2010). Issue 1.

[15] Pamela Stecker. 2009. *The Fashion Design Manual.* Palgrave Macmillan.

[16] Maryam Tavakol and Ulf Brefeld. 2014. Factored MDPs for Detecting Topics of User Sessions. In *Proceedings of the 8th ACM Conference on Recommender Systems (RecSys '14).* ACM, New York, NY, USA, 33–40. DOI:http://dx.doi.org/10.1145/2645710.2645739

[17] Bartlomiej Twardowski. 2016. Modelling Contextual Information in Session-Aware Recommender Systems with Neural Networks. In *Proceedings of the 10th ACM Conference on Recommender Systems (RecSys '16).* ACM, New York, NY, USA, 273–276. DOI:http://dx.doi.org/10.1145/2959100.2959162

[18] Manasi Vartak and Samuel Madden. 2013. CHIC: A Combination-based Recommendation System. In *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data (SIGMOD '13).* New York, NY, USA, 981–984.

[19] Jason Weston, Ron J. Weiss, and Hector Yee. 2013. Nonlinear Latent Factorization by Embedding Multiple User Interests. In *Proceedings of the 7th ACM Conference on Recommender Systems (RecSys '13).* ACM, New York, NY, USA, 65–68. DOI:http://dx.doi.org/10.1145/2507157.2507209

[20] Tong Zhao, Julian McAuley, and Irwin King. 2015. Improving Latent Factor Models via Personalized Feature Projection for One Class Recommendation. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management (CIKM '15).* ACM, New York, NY, USA, 821–830. DOI:http://dx.doi.org/10.1145/2806416.2806511